The Cosmic Data Archive

Salman Habib (T-8), Katrin Heitmann (ISR-1), Paul Ricker (University of Chicago at Urbana-Champaign), and Michael S. Warren (T-6); ***habib@lanl.gov***

As cosmological observations become richer in depth, coverage, resolution, and the number of observed quantities, it has become clear that predictive simulations will be necessary to fully extract all the information present in the data. Next-generation observations such as deep cluster surveys, supernova observations, and weak gravitational lensing surveys promise data with errors in the range of 1%. In order to make full use of this remarkable future dataset, simulations must be error-controlled to the subpercent level. The present state of the art needs to be improved by roughly an order of magnitude in the next several years. To reach this goal, a concerted community effort is required.

We have recently set up a new source for publicly available cosmological simulations—the Cosmic Data ArXiv (Fig. 1). The aim of this database is to provide scientists with state-of-the-art cosmological simulations to be used for a variety of purposes. The first set of data made available is from a code comparison study [1]. This dataset includes particle outputs at redshift $z = 0$ and initial conditions for different box-sizes. The results were generated with six different codes. Other researchers can test their own codes against the available data or contribute to the project by finding new methods to analyze and evaluate the simulation results.

Our simulations are employed in a variety of projects, ranging from Lyman-alpha forest analysis for the mass power spectrum in the Sloan Digital Sky Survey, the morphology of voids, studies of clusters and cluster distributions, determination of halo mass functions, etc. Due to the diversity of these topics we work with different kinds of codes: pure dark matter codes with high force and mass resolution, and codes including hydrodynamics implemented in simplified as well as sophisticated ways. We have worked on the development of three different codes, which we will briefly describe in the following text.

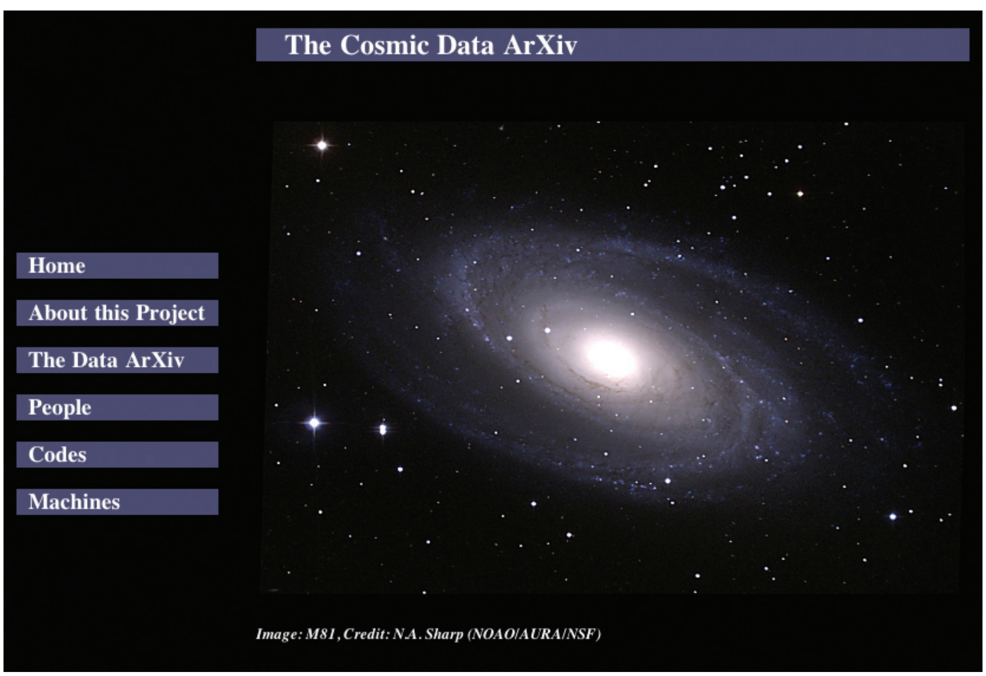

Figure 1—
The webpage for the Cosmic Data Arxiv: http://t8web.lanl.gov/people/heitmann/arxiv/.

The Mesh-based Cosmology Code MC^2 is a parallel particle mesh (PM) dark matter code developed at Los Alamos, which incorporates a simplified treatment of baryons via the Hydro-Particle-Mesh (HPM) method as well as neutrino modules. The code is designed to provide excellent performance for maximum values of the number of particles and the number of grid points: $N_p = N_g = 2048^3$; this number will increase as available computational resources continue to improve in size and performance. Because of the ubiquity of periodic boundary conditions in cosmology problems, MC^2 uses an FFT-based solver for the Poisson equation; time-stepping is handled via a symplectic method.

The Hashed Oct-Tree Code HOT is a dark matter tree-code also developed at Los Alamos National Laboratory. This code has defined the state of the art in high-resolution cosmological N-body simulations over the last decade. A smooth particle hydrodynamic (SPH) hydro capability for HOT has been developed and preliminary tests have been conducted successfully. The basic algorithm underlying the HOT code may be divided into several stages. First, particles are domain-decomposed into spatial groups. Second, a distributed tree data structure is constructed. In the main stage of the algorithm, this tree is traversed independently in each processor, with requests for nonlocal data being generated as needed. HOT has been run on a variety of parallel platforms over the last decade and has garnered Gordon Bell awards in 1992 (first place, performance category), and in 1997 (first place, price-performance category).

FLASH is an Adaptive Mesh Refinement (AMR) code for treating astrophysical hydrodynamics problems. It was originally developed at the Department of Energy Advanced Simulation and Computing (ASC) Alliances Center for Astrophysical Thermonuclear Flashes at the University of Chicago for the purpose of simulating Type Ia supernovae, novae, and x-ray bursts. It has since evolved to handle more general astrophysical problems, including those involving collisionless particle dynamics relevant for cosmological simulations. FLASH is freely available from the ASC Flash Center.

In the future, we expect the archive to naturally split into multiple subdomains covering areas such as initial conditions, raw code output, analysis routines and their results, error and convergence test suites, mock catalogs for different surveys, etc. The initial response to this initiative has been very enthusiastic; external contributions have already been received and have been added to the database.

[1] K. Heitmann, P.M. Ricker, M.S. Warren, and S. Habib, Los Alamos National Laboratory report LA-UR-04-5954 (2004), astro-ph/0411795 (submitted to *ApJS*).